

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:

- (A) NAME: MEDIGEN SA - c/o LENZ and STAHELIN
- (B) STREET: 25 Grand Rue
- (C) CITY: GENEVA
- (E) COUNTRY: SWITZERLAND
- (F) POSTAL CODE (ZIP): CH 1211

(ii) TITLE OF INVENTION: METHODS FOR DIAGNOSIS AND THERAPY OF AUTOIMMUNE DISEASE, SUCH AS INSULIN DEPENDENT DIABETES MELLITUS, INVOLVING RETROVIRAL SUPERANTIGENS

(iii) NUMBER OF SEQUENCES: 46

(iv) COMPUTER READABLE FORM:

- (A) MEDIUM TYPE: Floppy disk
- (B) COMPUTER: IBM PC compatible
- (C) OPERATING SYSTEM: PC-DOS/MS-DOS
- (D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

(v) CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/EP

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 25 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "DNA (synthetic)"

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION: 1..25
- (D) OTHER INFORMATION: /note= "page 11"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

TTTTTGAGTC CCCTTAGTAT TTATT

25

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "Synthetic DNA"

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION: 1..20
- (D) OTHER INFORMATION: /note= "page 26"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

ATCCAACAAAC CATGATGGAG

20

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "Synthetic DNA"

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION:1..21
- (D) OTHER INFORMATION:/note= "page 26"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

TCTCGTAAGG TGCAAATGAA G

21

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "synthetic DNA"

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION:1..21
- (D) OTHER INFORMATION:/note= "PAGE 26"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

GTAAAGGATC AAGTGCTGTG C

21

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "synthetic DNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

CTTTACAAAG CAGTATTGCT GC

22

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "synthetic DNA"

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION: 1..23
- (D) OTHER INFORMATION: /note= "RT 1a page 50"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

YAAATGGMGW AYGYTAACAG ACT

23

(2) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "Synthetic DNA"

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION: 1..23
- (D) OTHER INFORMATION: /note= "RT 1b page 50"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

YAAATGGMGW AYGYTAACTG ACT

23

(2) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 28 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "synthetic DNA"

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION:1..28
- (D) OTHER INFORMATION:/note= "RT 2a-nested page 50"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

CGTCTAGAGC CTCTCCGGCA TGATCCCG

28

(2) INFORMATION FOR SEQ ID NO: 9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 28 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "synthetic DNA"

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION:1..28
- (D) OTHER INFORMATION:/note= "RT 2b-nested page 50"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

CGTCTAGAGC CTCTCCGGCA TGATCCCA

28

(2) INFORMATION FOR SEQ ID NO: 10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "synthetic DNA"

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION:1..21
- (D) OTHER INFORMATION:/note= "Common 5' anchor page 50"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

TGCGCCAGCA ATGTATCCAT G

21

(2) INFORMATION FOR SEQ ID NO: 11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "SYNTHETIC DNA"

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION: 1..21
- (D) OTHER INFORMATION: /note= "1K1,2-1 page 50"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

GGGTGGCAGT GCATCATAGG T

21

(2) INFORMATION FOR SEQ ID NO: 12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "synthetic DNA"

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION: 1..22
- (D) OTHER INFORMATION: /note= "4K1,2-4 page 51"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

GGGAGAGGGT CAGCAGCAGA CA

22

(2) INFORMATION FOR SEQ ID NO: 13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "SYNTHETIC DNA"

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION:1..22
- (D) OTHER INFORMATION:/note= "K1,2-10 page 51"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

GACAGCAAGC CAGTGATAAG CA

22

(2) INFORMATION FOR SEQ ID NO: 14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 19 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "synthetic DNA"

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION:1..19
- (D) OTHER INFORMATION:/note= "K1,2-16 page 51"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

GGAACAGGGGA CTCTCTGCA

19

(2) INFORMATION FOR SEQ ID NO: 15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "synthetic DNA"

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION:1..20
- (D) OTHER INFORMATION:/note= "K1,2-17 page 51"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

GGGAAGGGTA AGGAAGTGTG

20

(2) INFORMATION FOR SEQ ID NO: 16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 19 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "synthetic DNA"

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION:1..19
- (D) OTHER INFORMATION:/note= "K1,2-22 page 51"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

GGTGTTCCTC CTGAGGGAG

19

(2) INFORMATION FOR SEQ ID NO: 17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "synthetic DNA"

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION:1..21
- (D) OTHER INFORMATION:/note= "K1,2-26 page 51"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

GAAGAATGGC CAACAGAAGC T

21

(2) INFORMATION FOR SEQ ID NO: 18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "synthetic DNA"

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION:1..20
- (D) OTHER INFORMATION:/note= "K1,2-27 page 51"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:

GGGAAACAAG GAGTGTGAGT

20

(2) INFORMATION FOR SEQ ID NO: 19:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 39 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "synthetic DNA"

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION:1..39
- (D) OTHER INFORMATION:/note= "U3-R-poly(AS) common page 51"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:

CATGTATATG CGGCCGCTGC GCCAGCAATG TATCCATGG

39

(2) INFORMATION FOR SEQ ID NO: 20:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "synthetic DNA"

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION:1..21
- (D) OTHER INFORMATION:/note= "RT1 page 51"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 20:

TATCTTCGT TTCTGCAGCA C

21

(2) INFORMATION FOR SEQ ID NO: 21:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "synthetic DNA"

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION:1..22
- (D) OTHER INFORMATION:/note= "RT2 page 51"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:

TAACTGGTTG AAGAGCTCGA CC

22

(2) INFORMATION FOR SEQ ID NO: 22:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "synthetic DNA"

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION:1..21
- (D) OTHER INFORMATION:/note= "R-U5-1 page 51"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 22:

ATACTAAGGG GACTCAGAGG C

21

(2) INFORMATION FOR SEQ ID NO: 23:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "synthetic DNA"

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION:1..27
- (D) OTHER INFORMATION:/note= "R-U5-2 page 51"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 23:

CAGAGGCTGG TGGGATCCTC CATATGC

27

(2) INFORMATION FOR SEQ ID NO: 24:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 25 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "SYNTHETIC DNA"

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION:1..25
- (D) OTHER INFORMATION:/note= "page 52"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 24:

TTTTTGAGTC CCCTTAGTAT TTATT

25

(2) INFORMATION FOR SEQ ID NO: 25:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "synthetic DNA"

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION:1..22
- (D) OTHER INFORMATION:/note= "Page 52"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 25:

AGGTATTGTC CAAGGTTCT CC

22

(2) INFORMATION FOR SEQ ID NO: 26:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 22 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: other nucleic acid
 - (A) DESCRIPTION: /desc = "SYNTHETIC DNA"
- (ix) FEATURE:
 - (A) NAME/KEY: misc_feature
 - (B) LOCATION:1..22
 - (D) OTHER INFORMATION:/note= "page 52"
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 26:

CTTTACAAAG CAGTATTGCT GC

22

(2) INFORMATION FOR SEQ ID NO: 27:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 21 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: other nucleic acid
 - (A) DESCRIPTION: /desc = "synthetic DNA"
- (ix) FEATURE:
 - (A) NAME/KEY: misc_feature
 - (B) LOCATION:1..21
 - (D) OTHER INFORMATION:/note= "page 52"
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 27:

GTAAAGGATC AAGTGCTGTG C

21

(2) INFORMATION FOR SEQ ID NO: 28:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 29 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: other nucleic acid
 - (A) DESCRIPTION: /desc = "synthetic DNA"

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION:1..29
- (D) OTHER INFORMATION:/note= "page 53"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 28:

GACTAAGCTT AAGAACCCAT CAGAGATGC

29

(2) INFORMATION FOR SEQ ID NO: 29:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 31 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "synthetic DNA"

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION:1..31
- (D) OTHER INFORMATION:/note= "page 53"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 29:

AGACTGGATC CGTTAAGTCG CTATCGACAG C

31

(2) INFORMATION FOR SEQ ID NO: 30:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 208 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "RETROVIRAL DNA"

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION:1..208
- (D) OTHER INFORMATION:/note= "FIGURE 7A"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 30:

CATCTCCCTC AGGAGAAACA CCCACGAATG ATCAATAAAAT ACTAAGGGGA CTCAGAGGCT	60
GGTGGGATCC TCCATATGCT GAACGTTGGT TCCCGGGGCC CCCTTATTTC TTTCTCTATA	120
CTTTGTCTCT GTGTCTTTT CTTTCCAAG TCTTCTTCAT TTGCACCTTA CGAGAAACAT	180
CTCCATCATG GTTGTGGAT GGGGGCAA	208

(2) INFORMATION FOR SEQ ID NO: 31:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1060 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "RETROVIRAL DNA"

(ix) FEATURE:

- (A) NAME/KEY: misc feature
- (B) LOCATION: 1..1060
- (D) OTHER INFORMATION: /note= "FIGURE 7B"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 31:

CTGCAGGTGT ACCCAACAGC TCCGAAGAGA CAGTGACATC GAGAACGGGC CATGATGACG	60
ATGGCGGTTT TGTCGAAAAG AAAAGGGGA AATGTGGGG AAGCAAGAG AGATGAGATT	120
GTTACTGTGT CTGTATAGAA AGAAGTAGAC ATAGGAGACT CCATTTGTT CTGTACTAAG	180
AAAAATTCTT CTGCCTTGAG ATGCTGTTAA TCTATGACCT TACCCCCAAC CCCGTGCTCT	240
CTGAAACATG TGCCGTGTCA AACTCAGGGT TAAATGGATT AAGGGTGGTG CAAGATGTGC	300
TTTGTAAAC AGATGCTTGA AGGCAGCATG CTCATTAAGA GTCATCACCA CTCCCTAATC	360
TCAAGTACCC AGGGACACAA ACAC TGCAGA AGGCCGCAGG GACCTCTGCC TAGGAAAGCC	420
AGGTATTGTC CAAGGTTTCT CCCCATGTGA TAGTCTGAAA TATGGCCTCG TGGGAAGGGA	480
AAGACCTGAC CATCCCCAG ACCAACACCC GTAAAGGTC TGTGCTGAGG AGGATTAGTA	540
TAAGAGGAAA GCATGCCTCT TGCAGTTGAG AGAAGAGGAA GACATCTGTC TCCTGCCAT	600
CCCCTGGCA ATGGAAATGTC TCAGTATAAA ACCCGATTGA ACATTCCATC TACTGAGATA	660
GGGAAAAACT GCCTTAGGGC TGGAGGTGGG ACATGTGGGC AGCAATACTG CTTTGTAAAG	720
CATTGAGATG TTTATGTGTA TGTATATCTA AAAGCACAGC ACTTGATCCT TTACCTTGTC	780
TATGATGCAA ACACCTTTGT TCACGTGTTT GTCTGCTGAC CCTCTCCCCA CTATTGTCTT	840
GTGACCCTGA CACATCTCCC TCAGGAGAAA CACCCACGAA TGATCAATAA ATACTAAGGG	900

GAATCAGAGG CTGGTGGGAT CCTCCATATG CTGAACGTTG GTTCCCAGGG CCCCTTATT	960
TCTTTCTCTA TACTTTGTCT CTGTGTCTTT TTCTTTCCA AGTCCTCTTC ATTTGCACCT	1020
TACGAGAAC ATCTCCATCA TGGTTGTTGG ATGGGGGCAA	1060

(2) INFORMATION FOR SEQ ID NO: 32:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1754 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "RETROVIRAL DNA"

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION: 1..1754
- (D) OTHER INFORMATION: /note = "FIGURE 7C"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 32:

ATGGTAAACAC CAGTCACATG GATGGATAAT CCTATAGAAAG TATATGTTAA TGATAGTGTA	60
TGGGTACCTG GCCCCCACAGA TGATCGCTGC CCTGCCAAC CTGAGGAAGA AGGGATGATG	120
ATAAAATATT CCATTGGGTA TCATTATCCT CCTATTTGCC TAGGGAGAGC ACCAGGATGT	180
TTAATGCCTG CAGTCCAAAA TTGGTTGGTA GAAGTACCTA CTGTCAGTCC TAACAGTAGA	240
TTCACTTATC ACATGGTAAG CGGGATGTCA CTCAGGCCAC GGGTAAATTAA TTTACAAGAC	300
TTTTCTTATC AAAGATCATT AAAATTAGA CCTAAAGGGA AAACCTGCC CAAGGAAATT	360
CCTAAAGGAT CAAAGAATAC AGAAGTTTA GTTGGGAAG AATGTGTGGC CAATAGTGTG	420
GTGATATTAC AAAACAATGA ATTGGAACT ATTATAGATT AGGCACCTCG AGGTCAATT	480
TACCACAATT GCTCAGGACA AACTCAGTCG TGTCCAAGTG CACAAGTGAG TCCAGCTGTC	540
GATAGCGACT TAACAGAAAG TCTAGACAAA CATAAGCATA AAAAATTACA GTCTTTCTAC	600
CTTTGGGAAT GGGAGAAAA AGGAATCTCT ACCCCAAAGAC CAAAAATAAT AAGTCCTGTT	660
TCTGGTCTG AACATCCAGA ATTGTGGAGG CTTACTGTGG CCTCACACCA CATTAGAATT	720
TGGTCTGGAA ATCAAACATT AGAAACAAGA TATCGTAAGC CATTATAC TATCGACCTA	780
AATTCCATTC TAACGGTTCC TTTACAAAGT TGCCTAAAGC CCCCTTATAT GCTAGTTGTA	840
GGAAATATAG TTATTAACC AGCCTCCAA ACTATAACCT GTGAAAATTG TAGATTGTTT	900
ACTTGCATTG ATTCAACTTT TAATTGGCAG CACCGTATTG TGCTGGTGAG AGCAAGAGAA	960
GGCATGTGGA TCCCTGTGTC CACGGACCGA CGTGGGAGG CCTCGCCATC CATCCATATT	1020

TTGACTGAAA TATTAAGG CGTTAAAT AGATCCAAA GATTCACTT TACTTTAATT	1080
GCAGTGATTA TGGGATTAAAT TGCAGTCACA GCTACGGCTG CTGTGGCAGG GGTTGCATTG	1140
CACTCTCTG TTCAGTCAGT AAACTTGTT AATTATTGGC AAAAGAATTG TACAAGATTG	1200
TGGAATTACAC AATCTAGTAT TGATCAAAA TTGGCAAGTC AAATTAATGA TCTTAGACAA	1260
ACTGTCACTT GGATGGGAGA CAGGCTTGAC TTAGAACATC ATTTCCAGTT ACAGTGTGAC	1320
TGGAATACGT CAGATTTTG TATTACACCC CAAATTTATA ATGAGTCTGA GCATCACTGG	1380
GACATGGTTA GACGCCATCT ACAGGGAAAGA GAAGATAATC TCACCTTACA CATTCCAAA	1440
TTAAAAGAAC AAATTTCGA AGCATCAAAA GCCCATTAA ATTTGGTGCC AGGAACGTGAG	1500
GCAATTGCAG GAGTTGCTGA TGGCCTCGCA AATCTAACCT CTGTCACTTG GATTAAGACC	1560
ATCAGAAGTA CTATGATTAT AAATCTCATA TTAATCGTTG TGTGCCTGTT TTGTCTGTTG	1620
TTAGTCTGCA GGTGTACCCC AACAGCTCCG AAAAAAACAG TGACATCGAG AACGGGCCAT	1680
GAATGACAAA GGCGGTTTTT GTTCAAAAAA AAAAGGGGG AAATTTGGG GAAAACCAAA	1740
AAAATGAAAA TGTT	1754

(2) INFORMATION FOR SEQ ID NO: 33:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 520 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: other nucleic acid
 - (A) DESCRIPTION: /desc = "RETROVIRAL DNA"
- (ix) FEATURE:
 - (A) NAME/KEY: misc_feature
 - (B) LOCATION: 1..520
 - (D) OTHER INFORMATION: /note= "FIGURE 7D"
- (ix) FEATURE:
 - (A) NAME/KEY: CDS
 - (B) LOCATION: 59..517

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 33:

ACATTTGAAG TTCTACAATG AACCCATCAAG AGATGCAAAG AAAAGCGCCT CCACGGAG	58
ATG GTA ACA CCA GTC ACA TGG ATG GAT AAT CCT ATA GAA GTA TAT GTT	106
Met Val Thr Pro Val Thr Trp Met Asp Asn Pro Ile Glu Val Tyr Val	
1 5 10 15	
AAT GAT AGT GTA TGG GTA CCT GGC CCC ACA GAT GAT CGC TGC CCT GCC	154
Asn Asp Ser Val Trp Val Pro Gly Pro Thr Asp Asp Arg Cys Pro Ala	
20 25 30	

AAA CCT GAG GAA GAA GGG ATG ATG ATA AAT ATT TCC ATT GGG TAT CAT	202
Lys Pro Glu Glu Glu Gly Met Met Ile Asn Ile Ser Ile Gly Tyr His	
35 40 45	
TAT CCT CCT ATT TGC CTA GGG AGA GCA CCA GGA TGT TTA ATG CCT GCA	250
Tyr Pro Pro Ile Cys Leu Gly Arg Ala Pro Gly Cys Leu Met Pro Ala	
50 55 60	
GTC CAA AAT TGG TTG GTA GAA GTA CCT ACT GTC AGT CCT AAC AGT AGA	298
Val Gln Asn Trp Leu Val Glu Val Pro Thr Val Ser Pro Asn Ser Arg	
65 70 75 80	
TTC ACT TAT CAC ATG GTA AGC GGG ATG TCA CTC AGG CCA CGG GTA AAT	346
Phe Thr Tyr His Met Val Ser Gly Met Ser Leu Arg Pro Arg Val Asn	
85 90 95	
TAT TTA CAA GAC TTT TCT TAT CAA AGA TCA TTA AAA TTT AGA CCT AAA	394
Tyr Leu Gln Asp Phe Ser Tyr Gln Arg Ser Leu Lys Phe Arg Pro Lys	
100 105 110	
GGG AAA ACT TGC CCC AAG GAA ATT CCT AAA GGA TCA AAG AAT ACA GAA	442
Gly Lys Thr Cys Pro Lys Glu Ile Pro Lys Gly Ser Lys Asn Thr Glu	
115 120 125	
GTT TTA GTT TGG GAA GAA TGT GTG GCC AAT AGT GTG GTG ATA TTA CAA	490
Val Leu Val Trp Glu Glu Cys Val Ala Asn Ser Val Val Ile Leu Gln	
130 135 140	
AAC AAT GAA TTC GGA ACT ATT ATA GAT TAG	520
Asn Asn Glu Phe Gly Thr Ile Ile Asp	
145 150	

(2) INFORMATION FOR SEQ ID NO: 34:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 153 amino acids
- (B) TYPE: amino acid
- (C) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 34:

Met Val Thr Pro Val Thr Trp Met Asp Asn Pro Ile Glu Val Tyr Val	
1 5 10 15	
Asn Asp Ser Val Trp Val Pro Gly Pro Thr Asp Asp Arg Cys Pro Ala	
20 25 30	
Lys Pro Glu Glu Glu Gly Met Met Ile Asn Ile Ser Ile Gly Tyr His	
35 40 45	
Tyr Pro Pro Ile Cys Leu Gly Arg Ala Pro Gly Cys Leu Met Pro Ala	
50 55 60	
Val Gln Asn Trp Leu Val Glu Val Pro Thr Val Ser Pro Asn Ser Arg	
65 70 75 80	
Phe Thr Tyr His Met Val Ser Gly Met Ser Leu Arg Pro Arg Val Asn	
85 90 95	

Tyr Leu Gln Asp Phe Ser Tyr Gln Arg Ser Leu Lys Phe Arg Pro Lys
 100 105 110
 Gly Lys Thr Cys Pro Lys Glu Ile Pro Lys Gly Ser Lys Asn Thr Glu
 115 120 125
 Val Leu Val Trp Glu Glu Cys Val Ala Asn Ser Val Val Ile Leu Gln
 130 135 140
 Asn Asn Glu Phe Gly Thr Ile Ile Asp
 145 150

(2) INFORMATION FOR SEQ ID NO: 35:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 603 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: other nucleic acid
 - (A) DESCRIPTION: /desc = "retroviral DNA"
- (ix) FEATURE:
 - (A) NAME/KEY: misc_feature
 - (B) LOCATION: 1..603
 - (D) OTHER INFORMATION:/note= "FIGURE 7E"
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 35:

ACATTTGAAG TTCTACAATG AACCCATCAG AGATGCAAAG AAAAGCGCCT CCACGGAGAT	60
GGTAACACCA GTCACATGGA TGGATAATCC TATAGAAGTA TATGTTAATG ATAGTGTATG	120
GGTACCTGGC CCCACAGATG ATCGCTGCC TGCCAAACCT GAGGAAGAAG GGATGATGAT	180
AAATATTTCC ATTGGGTATC ATTATCCTCC TATTTGCCTA GGGAGAGCAC CAGGATGTTT	240
AATGCCTGCA GTCCAAAATT GGTTGGTAGA AGTACCTACT GTCAGTCCTA ACAGTAGATT	300
CACTTATCAC ATGGTAAGCG GGATGTCACT CAGGCCACGG GTAAATTATT TACAAGACTT	360
TTCTTATCAA AGATCATTAA AATTTAGACC TAAAGGGAAA ACTTGCCCCA AGGAAATTCC	420
TAAAGGATCA AAGAATACAG AAGTTTTAGT TTGGGAAGAA TGTGTGGCCA ATAGTGTGGT	480
GATATTACAA AACAAATGAAT TCGGAACATAT TATAGATTAG GCACCTCGAG GTCAATTCTA	540
CCACAATTGC TCAGGACAAA CTCAGTCGTG TCCAAGTGCA CAAGTGAGTC CAGCTGTCGA	600
TAG	603

(2) INFORMATION FOR SEQ ID NO: 36:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 561 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(ix) FEATURE:

- (A) NAME/KEY: Protein
- (B) LOCATION: 1..561
- (D) OTHER INFORMATION: /note= "Figure 7F"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 36:

Met	Val	Thr	Pro	Val	Thr	Trp	Met	Asp	Asn	Pro	Ile	Glu	Val	Tyr	Val
1				5							10			15	

Asn	Asp	Ser	Val	Trp	Val	Pro	Gly	Pro	Thr	Asp	Asp	Arg	Cys	Pro	Ala
			20			25			25			30			

Lys	Pro	Glu	Glu	Glu	Gly	Met	Met	Ile	Asn	Ile	Ser	Ile	Gly	Tyr	His
		35				40						45			

Tyr	Pro	Pro	Ile	Cys	Leu	Gly	Arg	Ala	Pro	Gly	Cys	Leu	Met	Pro	Ala
			50			55						60			

Val	Gln	Asn	Trp	Leu	Val	Glu	Val	Pro	Thr	Val	Ser	Pro	Asn	Ser	Arg
65				70					75				80		

Phe	Thr	Tyr	His	Met	Val	Ser	Gly	Met	Ser	Leu	Arg	Pro	Arg	Val	Asn
			85			90			95						

Tyr	Leu	Gln	Asp	Phe	Ser	Tyr	Gln	Arg	Ser	Leu	Lys	Phe	Arg	Pro	Lys
		100					105				110				

Gly	Lys	Thr	Cys	Pro	Lys	Glu	Ile	Pro	Lys	Gly	Ser	Lys	Asn	Thr	Glu
		115				120						125			

Val	Leu	Val	Trp	Glu	Glu	Cys	Val	Ala	Asn	Ser	Val	Val	Ile	Leu	Gln
130				135							140				

Asn	Asn	Glu	Phe	Gly	Thr	Ile	Ile	Asp	Glx	Ala	Pro	Arg	Gly	Gln	Phe
145					150				155			160			

Tyr	His	Asn	Cys	Ser	Gly	Gln	Thr	Gln	Ser	Cys	Pro	Ser	Ala	Gln	Val
			165				170				175				

Ser	Pro	Ala	Val	Asp	Ser	Asp	Leu	Thr	Glu	Ser	Leu	Asp	Lys	His	Lys
			180				185				190				

His	Lys	Lys	Leu	Gln	Ser	Phe	Tyr	Leu	Trp	Glu	Trp	Glu	Glu	Lys	Gly
			195			200			205						

Ile	Ser	Thr	Pro	Arg	Pro	Lys	Ile	Ile	Ser	Pro	Val	Ser	Gly	Pro	Glu
210				215						220					

His	Pro	Glu	Leu	Trp	Arg	Leu	Thr	Val	Ala	Ser	His	His	Ile	Arg	Ile
225				230				235			240				

Trp Ser Gly Asn Gln Thr Leu Glu Thr Arg Tyr Arg Lys Pro Phe Tyr
 245 250 255
 Thr Ile Asp Leu Asn Ser Ile Leu Thr Val Pro Leu Gln Ser Cys Leu
 260 265 270
 Lys Pro Pro Tyr Met Leu Val Val Gly Asn Ile Val Ile Lys Pro Ala
 275 280 285
 Ser Gln Thr Ile Thr Cys Glu Asn Cys Arg Leu Phe Thr Cys Ile Asp
 290 295 300
 Ser Thr Phe Asn Trp Gln His Arg Ile Leu Leu Val Arg Ala Arg Glu
 305 310 315 320
 Gly Met Trp Ile Pro Val Ser Thr Asp Arg Pro Trp Glu Ala Ser Pro
 325 330 335
 Ser Ile His Ile Leu Thr Glu Ile Leu Lys Gly Val Leu Asn Arg Ser
 340 345 350
 Lys Arg Phe Ile Phe Thr Leu Ile Ala Val Ile Met Gly Leu Ile Ala
 355 360 365
 Val Thr Ala Thr Ala Ala Val Ala Gly Val Ala Leu His Ser Ser Val
 370 375 380
 Gln Ser Val Asn Phe Val Asn Tyr Trp Gln Lys Asn Ser Thr Arg Leu
 385 390 395 400
 Trp Asn Ser Gln Ser Ser Ile Asp Gln Lys Leu Ala Ser Gln Ile Asn
 405 410 415
 Asp Leu Arg Gln Thr Val Ile Trp Met Gly Asp Arg Leu Asp Leu Glu
 420 425 430
 His His Phe Gln Leu Gln Cys Asp Trp Asn Thr Ser Asp Phe Cys Ile
 435 440 445
 Thr Pro Gln Ile Tyr Asn Glu Ser Glu His His Trp Asp Met Val Arg
 450 455 460
 Arg His Leu Gln Gly Arg Glu Asp Asn Leu Thr Leu Asp Ile Ser Lys
 465 470 475 480
 Leu Lys Glu Gln Ile Phe Glu Ala Ser Lys Ala His Leu Asn Leu Val
 485 490 495
 Pro Gly Thr Glu Ala Ile Ala Gly Val Ala Asp Gly Leu Ala Asn Leu
 500 505 510
 Asn Pro Val Thr Trp Ile Lys Thr Ile Arg Ser Thr Met Ile Ile Asn
 515 520 525
 Leu Ile Leu Ile Val Val Cys Leu Phe Cys Leu Leu Leu Val Cys Arg
 530 535 540
 Cys Thr Pro Thr Ala Pro Lys Lys Thr Val Thr Ser Arg Thr Gly His
 545 550 555 560
 Glu

(2) INFORMATION FOR SEQ ID NO: 37:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 604 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION:1..604
- (D) OTHER INFORMATION:/note= "FIGURE 7G"

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION:59..601

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 37:

ACATTTGAAG TTCTACAATG AACCCATCAG AGATGCAAAG AAAAGCGCCT CCACGGAG	58
ATG GTA ACA CCA GTC ACA TGG ATG GAT AAT CCT ATA GAA GTA TAT GTT Met Val Thr Pro Val Thr Trp Met Asp Asn Pro Ile Glu Val Tyr Val	106
1 5 10 15	
AAT GAT AGT GTA TGG GTA CCT GGC CCC ACA GAT GAT CGC TGC CCT GCC Asn Asp Ser Val Trp Val Pro Gly Pro Thr Asp Asp Arg Cys Pro Ala	154
20 25 30	
AAA CCT GAG GAA GAA GGG ATG ATG ATA AAT ATT TCC ATT GGG TAT CAT Lys Pro Glu Glu Gly Met Met Ile Asn Ile Ser Ile Gly Tyr His	202
35 40 45	
TAT CCT CCT ATT TGC CTA GGG AGA GCA CCA GGA TGT TTA ATG CCT GCA Tyr Pro Pro Ile Cys Leu Gly Arg Ala Pro Gly Cys Leu Met Pro Ala	250
50 55 60	
GTC CAA AAT TGG TTG GTA GAA GTA CCT ACT GTC AGT CCT AAC AGT AGA Val Gln Asn Trp Leu Val Glu Val Pro Thr Val Ser Pro Asn Ser Arg	298
65 70 75 80	
TTC ACT TAT CAC ATG GTA AGC GGG ATG TCA CTC AGG CCA CGG GTA AAT Phe Thr Tyr His Met Val Ser Gly Met Ser Leu Arg Pro Arg Val Asn	346
85 90 95	
TAT TTA CAA GAC TTT TCT TAT CAA AGA TCA TTA AAA TTT AGA CCT AAA Tyr Leu Gln Asp Phe Ser Tyr Gln Arg Ser Leu Lys Phe Arg Pro Lys	394
100 105 110	
GGG AAA ACT TGC CCC AAG GAA ATT CCT AAA GGA TCA AAG AAT ACA GAA Gly Lys Thr Cys Pro Lys Glu Ile Pro Lys Gly Ser Lys Asn Thr Glu	442
115 120 125	
GTT TTA GTT TGG GAA GAA TGT GTG GCC AAT AGT GTG GTG ATA TTA CAA Val Leu Val Trp Glu Glu Cys Val Ala Asn Ser Val Val Ile Leu Gln	490
130 135 140	

AAC AAT GAA TTC GGA ACT ATT ATA GAT TTA GGC ACC TCG AGG TCA ATT Asn Asn Glu Phe Gly Thr Ile Ile Asp Leu Gly Thr Ser Arg Ser Ile 145 150 155 160	538
CTA CCA CAA TTG CTC AGG ACA AAC TCA GTC GTG TCC AAG TGC ACA AGT Leu Pro Gln Leu Leu Arg Thr Asn Ser Val Val Ser Lys Cys Thr Ser 165 170 175	586
GAG TCC AGC TGT CGA TAG Glu Ser Ser Cys Arg 180	604

(2) INFORMATION FOR SEQ ID NO: 38:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 181 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 38:

Met Val Thr Pro Val Thr Trp Met Asp Asn Pro Ile Glu Val Tyr Val 1 5 10 15
Asn Asp Ser Val Trp Val Pro Gly Pro Thr Asp Asp Arg Cys Pro Ala 20 25 30
Lys Pro Glu Glu Glu Gly Met Met Ile Asn Ile Ser Ile Gly Tyr His 35 40 45
Tyr Pro Pro Ile Cys Leu Gly Arg Ala Pro Gly Cys Leu Met Pro Ala 50 55 60
Val Gln Asn Trp Leu Val Glu Val Pro Thr Val Ser Pro Asn Ser Arg 65 70 75 80
Phe Thr Tyr His Met Val Ser Gly Met Ser Leu Arg Pro Arg Val Asn 85 90 95
Tyr Leu Gln Asp Phe Ser Tyr Gln Arg Ser Leu Lys Phe Arg Pro Lys 100 105 110
Gly Lys Thr Cys Pro Lys Glu Ile Pro Lys Gly Ser Lys Asn Thr Glu 115 120 125
Val Leu Val Trp Glu Glu Cys Val Ala Asn Ser Val Val Ile Leu Gln 130 135 140
Asn Asn Glu Phe Gly Thr Ile Ile Asp Leu Gly Thr Ser Arg Ser Ile 145 150 155 160
Leu Pro Gln Leu Leu Arg Thr Asn Ser Val Val Ser Lys Cys Thr Ser 165 170 175
Glu Ser Ser Cys Arg 180

(2) INFORMATION FOR SEQ ID NO: 39:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 182 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(ix) FEATURE:

- (A) NAME/KEY: Protein
- (B) LOCATION: 1..182
- (D) OTHER INFORMATION:/note= "FIGURE 7H"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 39:

Phe	Thr	Ile	Pro	Leu	Ala	Glu	Gln	Asp	Cys	Glu	Lys	Phe	Ala	Phe	Thr
1				5					10				15		
Ile	Pro	Ala	Ile	Asn	Asn	Lys	Glu	Pro	Ala	Thr	Arg	Phe	Gln	Trp	Lys
			20				25					30			
Val	Leu	Pro	Gln	Gly	Met	Leu	Asn	Ser	Pro	Thr	Ile	Cys	Gln	Thr	Phe
	35				40			40			45				
Val	Gly	Arg	Ala	Leu	Gln	Pro	Val	Arg	Asp	Lys	Phe	Ser	Asp	Cys	Tyr
	50				55			55			60				
Ile	Ile	His	Tyr	Phe	Asp	Asp	Ile	Leu	Cys	Ala	Ala	Glu	Thr	Lys	Asp
	65				70			70		75		75		80	
Lys	Leu	Ile	Asp	Cys	Tyr	Thr	Phe	Leu	Pro	Ala	Glu	Val	Ala	Asn	Ala
			85				85		90			90		95	
Gly	Leu	Ala	Ile	Ala	Ser	Asp	Lys	Ile	Gln	Thr	Ser	Thr	Pro	Phe	His
			100				100		105			105		110	
Tyr	Leu	Gly	Met	Gln	Ile	Glu	Asn	Arg	Lys	Ile	Lys	Pro	Gln	Lys	Ile
	115					120			120			125			125
Glu	Ile	Arg	Lys	Asp	Thr	Leu	Lys	Thr	Leu	Asn	Asp	Phe	Gln	Lys	Leu
	130					135			135			140			
Leu	Gly	Asp	Ile	Asn	Trp	Ile	Arg	Pro	Thr	Leu	Gly	Ile	Pro	Thr	Tyr
	145					150			150		155		155		160
Ala	Met	Ser	Asn	Leu	Phe	Ser	Ile	Leu	Arg	Gly	Asp	Ser	Asp	Leu	Asn
			165				165		170			170		175	
Ser	Lys	Arg	Met	Leu	Thr										
			180												

(2) INFORMATION FOR SEQ ID NO: 40:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 250 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
 (A) DESCRIPTION: /desc = "retroviral DNA"

(ix) FEATURE:

(A) NAME/KEY: misc feature
 (B) LOCATION: 1..250
 (D) OTHER INFORMATION:/note= "FIGURE 8A"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 40:

GTAAATGACA CCTATGATGC ACTGCCACCC TTTCACTGTT TCACCCCTGAA CATCTGCTTT	60
TTACATCTAA GTGATTGTAC CCAATAAAATA GTGTGGAGAC CAGAGCTCTG AGCCTTTGTC	120
AGCCTCCATT TTGCAACTGG TCCCCTGGCT CCCACCTTTA TGAACCTTTA ACCTGTCTTT	180
TCTCATTCCT TTGTCACCAT TGGACTTTGG GTACCCCTACG GGTGGTGTG AGGCTGTCAC	240
CGCACATTAA	250

(2) INFORMATION FOR SEQ ID NO: 41:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 203 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
 (A) DESCRIPTION: /desc = "RETROVIRAL DNA"

(ix) FEATURE:

(A) NAME/KEY: misc feature
 (B) LOCATION: 1..203
 (D) OTHER INFORMATION:/note= "FIGURE 8B"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 41:

GTAGTTAA TCTATAATCT ATAGAGACAA TGCTTATCAC TGGCTTGCTG TCAATAAAATA	60
TGTGGGTAAA TCTCTGTTCA AGACTCTCAG CTTTGAAGCT GTGAGACCCC TGATTTCCCA	120
CTCCACACCT CTATATTCT GTGTGTGTGT CTTAATTCC TCCAGTGTG CTGGGTTAGG	180
GTCTCCTCGA CGAGCTGTCG TGC	203

(2) INFORMATION FOR SEQ ID NO: 42:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 283 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
 (A) DESCRIPTION: /desc = "RETROVIRAL DNA"

(ix) FEATURE:
 (A) NAME/KEY: misc_feature
 (B) LOCATION: 1..283
 (D) OTHER INFORMATION:/note= "FIGURE 8C"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 42:

AACTCAGCTG CTGCACAGTG GTCGAGCCTC CAGAGCTCAT GCCATTGCAG TGGTCAGAGC	60
CTGGCCCTCC TCTTCCTGCA TAGAACCTGG ATTCAATCTG TAAGGTGGGA AGTGCAGCAG	120
CAGAGAACTC TGGCCTTGCA GAGAGTCCTT GTTCCCACCTT CACTTCCCTT TTCACCAAAT	180
AAAACCCCTGC TTTCACTCAT GCATCAAATT GTCTGTGAGC CTACATTTT GTGGCCATGG	240
GACAAGAACCA CCATCTTTAG CTGAGCTAGG GAAAAGTCCT GCA	283

(2) INFORMATION FOR SEQ ID NO: 43:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 245 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
 (A) DESCRIPTION: /desc = "retroviral DNA"

(ix) FEATURE:
 (A) NAME/KEY: misc_feature
 (B) LOCATION: 1..245
 (D) OTHER INFORMATION:/note= "FIGURE 8D"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 43:

GATGTGACCA CTGTGACCTA CCTACACTGG AGATGGCTCA CACTTCCTTA CCCTTCCCCCT	60
GCTGTACCAA TAAATAACAG CACAGCCTGA CATTGGAGC CATTACCGGT CTTTGTGACT	120
TGGTGGTAGT GGTATCCCCT AGGGCCCAGC TGTCTTTCT TTTATCTCTT TGTCTTGTGT	180
CTTTATTTCT ATGAGTCTCT CGTCTCCGCA CATGGGGAGA AAAACCCATA GACCCTGTAG	240
GGCTG	245

(2) INFORMATION FOR SEQ ID NO: 44:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 181 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
 (A) DESCRIPTION: /desc = "retroviral DNA"

(ix) FEATURE:

(A) NAME/KEY: misc_feature
 (B) LOCATION: 1..181
 (D) OTHER INFORMATION: /note= "FIGURE 8E"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 44:

CTCACAAAAA TAATAAAAGC TTCTGTTGGC CATTCTTCAG ATCTTCATCT CTTGTGAGGA	60
TCCCCCTGTA CATGTAAAAA TGTAATAAAA CTTGTATCCT TTCTCCTCTT AATCTGTCTT	120
GCATCAATAT CATTCCCTAGA CCCAGTCAGA GATGGGTGGA GGTGAGCCGT ACATTTCCCT	180
A	181

(2) INFORMATION FOR SEQ ID NO: 45:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 287 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "RETROVIRAL DNA"

(ix) FEATURE:

(A) NAME/KEY: misc_feature
 (B) LOCATION: 1..287
 (D) OTHER INFORMATION: /note= "FIGURE 8F"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 45:

CAGAGAACTC CAGCCAGCTG TGATGGAGCC TCAGGAAGTT CACAGTTGCA GCAGGAAGGA	60
GCCTGGCTGC TCCTCTTCCT GTGTGGAACC TGGGATTAGA ACAGGGCTGGC AGGAAGTGCT	120
TTAGCAGGGA CTCTGGCCTA CTCACACTCC TTGTTTCCCC CCTTTCTTCC TTTTCACTCA	180
ATAAAGCCCT GTCTTACTCA CCATTCAAAT TGTCTGTGAG CCTGAATTTT CATGGCTGTG	240
GGACAAAGAA CCCTATTTT AGCTGAACTA AGGAAAATTC CTGCAAA	287

(2) INFORMATION FOR SEQ ID NO: 46:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 264 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "RETROVIRAL DNA"

(ix) FEATURE:

- (A) NAME/KEY: misc_feature
- (B) LOCATION: 1..264
- (D) OTHER INFORMATION: /note= "FIGURE 8G"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 46:

GTGATTGTCT GCTGACCCCTC TCCCCACAAT TGTCTTGTGA CCCTGACACA TCCCCCTCTT	60
CGAGAAACAC CCGCGGATGA TCAATAAAATA TTAAGGGAAC TCAGAGGCTG GCAGGATCCT	120
CCATATGCTG AACGCTGGTT GCCCCGGGTC CCCTTCTTTC TTTCTCTATA CTTTGTCTCT	180
GTGTCTTTTT CTTTCCAAA TCTCTCGTCC CACCTTACGA GAAACACCCA CAGGTGTGTC	240
CGGGCAACCC AACGCCACAT AACA	264